

Characterization of integrated hepatitis B viral DNA cloned from a human hepatoma and the hepatoma-derived cell line PLC/PRF/5

(hepatocellular carcinoma/expression of gene for hepatitis B surface antigen/bacteriophage λ cloning/electron microscopy)

ANNE DEJEAN, CHRISTIAN BRECHOT, PIERRE TIOLLAIS, AND SIMON WAIN-HOBSON

Unité de Recombinaison et Expression Génétique, Institut National de la Santé et de la Recherche Médicale U.163, Centre National de la Recherche Scientifique LA 271, Institut Pasteur, 28 rue du Dr. Roux, 75724 Paris Cedex 15, France

Communicated by Baruch S. Blumberg, January 26, 1983

ABSTRACT Recombinant phage clones carrying integrated hepatitis B virus (HBV) DNA sequences have been isolated from two phage libraries made from human DNA of a hepatoma and a hepatoma-derived cell line. One clone from each library has been characterized both by restriction mapping and by electron microscopy. In one clone there is at least one complete and uninterrupted HBV genome, and in the other the HBV sequences are composed of two major subgenomic fragments inverted with respect to each other. The host-virus junctions are localized within the positions 1,700–2,600 base pairs on the physical map of the free viral genome. The pre-S/S (surface antigen gene) region is conserved between the two clones. The two clones do not have common cellular sequences nor do they contain cellular homologues to six retroviral oncogenes. For one clone, the hepatitis B surface antigen gene was found to be functional when introduced into mouse thymidine kinase-negative cells by transfection.

Integration of hepatitis B virus (HBV) DNA into human liver DNA has been shown in many cases of hepatocellular carcinoma (HCC, also called hepatoma) and chronic hepatitis and in a few cases of acute and fulminant hepatitis (1–5). In a similar manner, integrated viral DNA of the woodchuck hepatitis B virus (WHV), a HBV-like virus, has been described in some cases of HCC and chronic hepatitis (ref. 6; J. Summers, personal communication).

Little is known about either the organization of the integrated HBV sequences or the mechanism of viral integration. It has been shown by Southern blot-hybridization analyses of liver DNA that there were a limited number of bands corresponding to integrated HBV DNA and that some of these bands could be oligomers of the viral genome (1–4). However, a relationship between HBV DNA integration and the life cycle of the virus or the clinical course of HBV infection has not yet been established.

In the limited number of advanced HCC cases in which both the tumorous and nontumorous parts of the liver were analyzed by the Southern blot-hybridization procedure, only integrated HBV sequences (i.e., without free viral DNA) were found in both parts (2, 3). Immunofluorescence studies have shown that the hepatitis B surface antigen (HBsAg) is usually absent from the poorly differentiated tumorous hepatocytes, whereas HBsAg is frequently found in the nontumorous part of the liver (7). It therefore appears that whereas in nontumorous hepatocytes HBsAg can be synthesized from integrated HBV sequences, in tumorous hepatocytes such expression is absent despite the presence of integrated HBV sequences.

To investigate the organization of integrated HBV sequences, the flanking cellular DNA, and the control of HBsAg

gene expression in HCC, we have isolated and analyzed HBV-containing recombinant phage clones from two libraries. These libraries were constructed from the liver DNA of a patient who died of liver failure due to HCC (1) and from the HCC-derived cell line PLC/PRF/5 (8).

MATERIALS AND METHODS

DNA. Patient 53 (a West African Black from the Ivory Coast), from whose tumorous liver DNA was extracted, was serum positive for both HBsAg and antibody to the HBV core antigen (HBcAg). By direct immunofluorescence the tumor tissue was negative for HBsAg and the HBV HBcAg. No nontumorous tissue was available (1). The PLC/PRF/5 cell line (8) was a culture of the original cell line after passage twice through *nude* mice. The DNA from patient 53 carries at least five integrated HBV sequences (1), and PLC/PRF/5 DNA carries six (1).

Phages. λ L47.1, a *Bam*HI vector, was kindly provided by W. Loenen (9).

Bacteria. Strain LA101 (10) was used for phage infections. For *in vitro* packaging the strains BHB2688 and BHB2690 (11) were kindly provided by John Collins.

Enzymes. All restriction enzymes and phage T4 DNA ligase were purchased from New England BioLabs. DNA polymerase I and DNase I were obtained from Boehringer.

Blot-Hybridization. Southern blot-hybridizations (12) were as modified by Wahl *et al.* (13). Nick-translated probes were prepared as described by Weinstock *et al.* (14).

Cloning. In brief, cellular DNA was partially digested with *Mbo* I and fractionated on a 5–40% sucrose gradient. The 10- to 20-kilobase pair (kb) fractions were pooled and the DNA was precipitated. The phage arms were prepared by first ligating the phage upon itself and then digesting with *Bam*HI and *Xho* I (the latter cuts within the internal fragment). The digest was fractionated on a 5–40% sucrose gradient, the phage arms were pooled, and the DNA was precipitated. The arms were redigested with *Bam*HI and *Xho* I; the background titer was less than 0.01% of the intact phage DNA. Ligation was performed at a total DNA concentration of 250 μ g ml⁻¹ and at a weight ratio of arms to liver DNA of 2:1. Three hundred units of T4 DNA ligase was used for a total of 30 μ g of DNA, and the reaction mixture was incubated at 12°C for 24 hr. Ligation was verified by gel electrophoresis. *In vitro* packaging lysates were made according to the method of Ish-Horowitz and Burke (15). The resulting efficiency of recombinant DNA was of the order of 5×10^5 plaque-forming units μ g⁻¹ of vector DNA. The phage libraries were screened by using the Benton and Davies *in situ*

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: HBV, hepatitis B virus; WHV, woodchuck hepatitis virus; HBsAg, HBV surface antigen; HBcAg, HBV core antigen; HCC, hepatocellular carcinoma; kb, kilobase pair(s); bp, base pair(s).

FIG. 1. Restriction and physical maps of the inserts of clones λ IA22 (A) and λ A3 (B). Open bars denote HBV sequences; the arrowheads within denote the same orientation as the open reading frames carried by free viral DNA. Single lines denote the cellular sequences. Restriction sites are: R, *EcoRI*; B, *BamHI*; Bg, *Bgl* II; H, *HindIII*; X, *Xho* I; Sa, *Sac* I; K, *Kpn* I; Sp, *Sph* I. The arrows above each map represent the lengths of HBV DNA in kb deduced from the structures shown by electron microscopy. The bars below represent features established by blotting using specific probes: C, gene C sequences; S, gene S sequences; r, highly repetitive cellular sequences. The small fragment of unique DNA, H 0.6 at the left end of the restriction map of λ IA22, is also indicated. The 2.9-kb *EcoRI* fragment used to elucidate the relationship between the two inverted subgenomic fragments in λ A3 is shown beneath the restriction map (R 2.9).

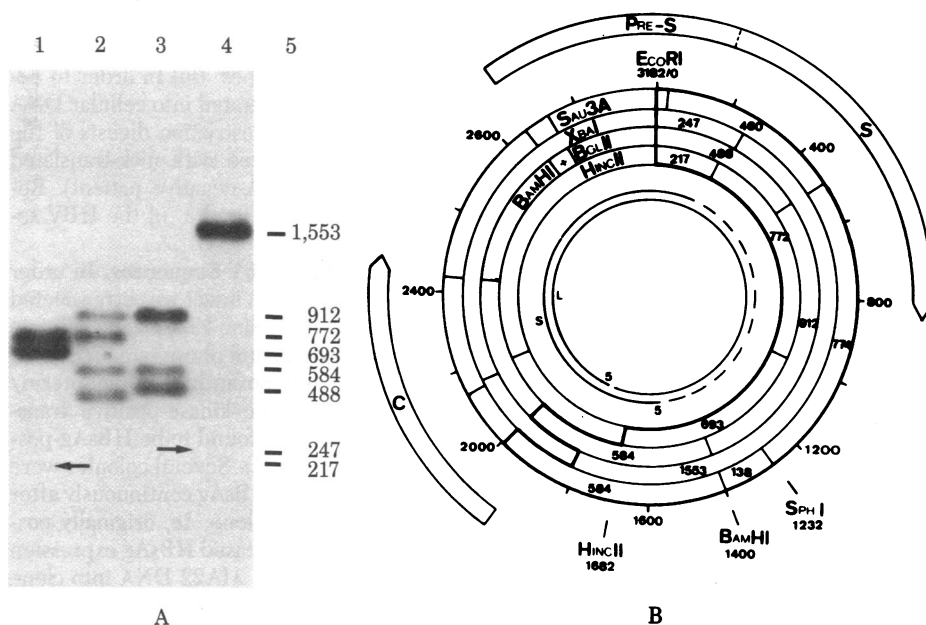


FIG. 2. (A) Autoradiograph of restriction digests of *Eco*RI-linearized cloned HBV (subtype ayw) DNA (22) blotted and hybridized with the 32 P-labeled nick-translated 2.9-kb *Eco*RI fragment of λ A3. Lane 1, *Hinc*II; lane 2, *Sau*3A; lane 3, *Bam*HI + *Bgl* II; lane 4, *Xba* I; lane 5, sizes of fragments in bp. The two small fragments, 247 and 217 bp, were visible on prolonged exposure. The uppermost band in lane 2 is due to an incomplete digestion. (B) Partial restriction map of HBV (subtype ayw) DNA. The HBV fragments hybridizing to the 2.9-kb *Eco*RI fragment, as shown in A, are surrounded by a bold line, and their sizes are given in bp. The open reading frames C, pre-S/S as well as the long (L) and short (S) strands of free viral DNA are given for orientation.

ref. 21). Such a site exists within the 2.9-kb *Eco*RI fragment at about 1.2 kb from the left *Eco*RI site. Thus by incorporating the data derived from Fig. 2 we concluded that it was the left subgenomic fragment that extended through to approximately position 1,550 bp.

Structure of Integrated Sequences: Electron Microscopy. The clones were characterized by heteroduplex analysis with M13 cloned single-stranded HBV DNA (23). Fig. 3 shows a typical heteroduplex between λ IA22 DNA and single-stranded HBV DNA. A smooth complete double-stranded circle of DNA of the same length as a HBV genome is clearly visible, as indicated by an arrow. The attached single-stranded circle is of M13 vector DNA, thereby localizing the *Eco*RI site within the double-

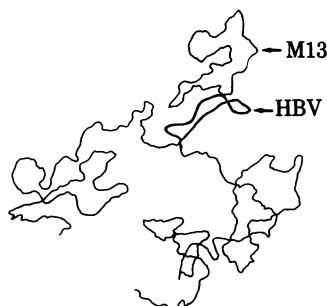
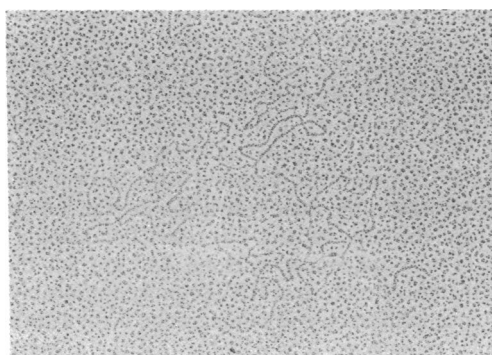


FIG. 3. Electron micrograph and drawing of heteroduplex between λ IA22 DNA and single-stranded HBV-M13 DNA. The double-stranded complete HBV genome is indicated in the drawing, as is the single-stranded DNA of the M13 vector.

stranded circle of DNA. The map determined from measuring 30 micrographs agreed well with the restriction map. More specifically, the distances between the *Eco*RI site and the origin of the two long single-stranded arms were 1.5 kb and 1.75 ± 0.1 kb. These sizes have been superimposed on the restriction map of λ IA22 (Fig. 1A). Only the shorter double-stranded region could be accommodated within the left-hand side of the HBV sequences because the *Eco*RI/*Eco*RI fragment at the left extremity of the HBV sequences is 1.55 kb. Thus the point of origin of the two long single-stranded arms on the double-stranded circle was close to, if not, the left cellular-HBV DNA junction. This site, about 1.45 kb left of the *Eco*RI site, was thus placed, by comparison of the map of λ IA22 with those of the three cloned HBV genomes of known sequence, within the cohesive end structure of the free viral DNA at position $1,750 \pm 100$ bp. Using a large molar excess single-stranded HBV DNA, we did not succeed in reproducibly finding other structures, including two M13-HBV molecules.

For clone λ A3, "snap-back" structures were found in preference to heteroduplexes between recombinant phage DNA and cloned HBV DNA. A typical stem-and-loop structure is shown in Fig. 4. The double-stranded stem is approximately 1.7 kb in length, and the loop is about 0.8 kb of single-stranded DNA. The map determined from measuring 20 micrographs placed this stem-and-loop structure within the HBV sequences located by restriction mapping. Even with a large molar excess of cloned HBV DNA, no structures between the loop and HBV DNA were reproducibly found. This stem-and-loop structure could be superimposed on the restriction map of λ A3 as follows (Fig. 1B): The center of the loop must be equidistant with respect to the *Eco*RI sites within the HBV sequences. The base of the stem was thus $1.7 + 0.8/2 = 2.1$ kb from the midpoint of the loop. Because the *Eco*RI sites are 2.9 kb apart, the base of the stem could be located at $2.1 - 2.9/2 \approx 0.6$ kb distant from the two *Eco*RI sites within the HBV sequences. For the left cellular junction this agreed with fine restriction mapping of this region (data not shown) and therefore the base of the stem probably corresponded to this junction. By comparison of the restriction map of the left subgenomic fragment with the complete maps of the three cloned HBV genomes, this site could be located between the C and pre-S coding regions at approximately position 2,600 kb (Fig. 2B). The top of the stem was at $1.7 - 0.6 = 1.1$ kb within the 2.9-kb *Eco*RI fragment. Thus the

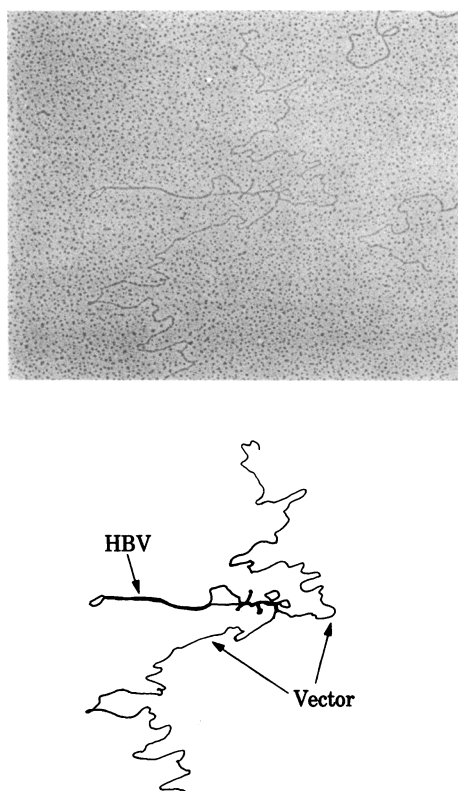


FIG. 4. Electron micrograph and drawing of a "snap-back" structure of λ A3 DNA. The stem-and-loop structure carrying HBV DNA is indicated. The small double-stranded regions due to inverted repeats in the liver cellular DNA are also clearly visible.

internal boundary of the right subgenomic fragment was at position 1,100 bp. For the right-hand side of the HBV sequences, the base of the stem was within the HBV subgenomic fragment because restriction mapping places HBV sequences beyond this position.

Also shown in the electron micrographs of clone λ A3 (Fig. 4) are small double-stranded regions of DNA due to inverted repeats within the flanking host cellular sequences (the phage vector alone does not form any such structures when similarly analyzed).

Cellular Flanking Sequences. For both clones the nature of the cellular sequences flanking the integrated HBV was analyzed: (i) The finding of discrete HBV hybridizing bands detected by Southern blots of genomic liver DNA (1–4) raised the question of the existence of a restricted number of integration sites within and amongst individuals. We looked for the presence of common flanking sequences amongst the clones derived from the tumor and the PLC/PRF/5 cell line. The small 0.6-kb *Hind*III fragment at the left-hand side of the insert in λ IA22 (H 0.6, Fig. 1A), which did not contain any repetitive DNA, was subcloned in the *Hind*III site of pBR322. The recombinant plasmid, p λ H6, was 32 P-labeled by nick-translation and used as a probe against filters carrying the 12 PLC/PRF/5 clones and the 2 other clones from patient 53. None of these clones hybridized with the λ IA22 flanking sequence. (ii) Because the two phages λ IA22 and λ A3 were cloned from HCC DNA and a HCC-derived cell line, respectively, we looked for the presence of cellular homologues to several retroviral oncogenes in the flanking sequences. The 3 clones derived from the tumor library and the 12 clones from the PLC/PRF/5 library were screened with the six retroviral oncogene-specific probes *v-abl*, *v-fps*, *v-ras*, *v-myc*, *v-myb*, and *v-src*. Under

nonstringent hybridization and washing conditions (0.9 M NaCl/0.09 M sodium citrate/50°C), none of the clones cross-hybridized with any of the six oncogene probes. (iii) In order to determine whether or not HBV was integrated into cellular DNA carrying highly repeated sequences, restriction digests of the two clones were blotted and hybridized with nick-translated total human DNA (from a HBV DNA-negative patient). Repetitive sequences were found both 5' and 3' of the HBV sequences as shown in Fig. 1A and B.

Expression of λ IA22 Integrated HBV Sequences. In order to ascertain whether the HBsAg gene is intact we cotransfected λ IA22 DNA along with a cloned thymidine kinase gene (17) into mouse Ltk⁻ cells by the modified calcium phosphate procedure (18). After 15 days of selection on hypoxanthine/aminopterin/thymidine medium for the thymidine kinase-positive transformants, the culture supernatant was found to be HBsAg-positive by radioimmunoassay (AUSRIA II). Several colonies were cloned. One clone, 2a, has expressed HBsAg continuously after passaging at least six times. Another clone, 1c, originally positive for HBsAg, proved unstable and ceased HBsAg expression after two passages. The transfection of λ IA22 DNA into clone 2a cellular DNA occurred without rearrangement of the cloned insert (data not shown). No free viral DNA was detected in the culture medium or in the cytoplasm.

After CsCl density gradient purification HBsAg activity resided in fractions of density 1.2 g ml⁻¹, typical of human HBsAg 22-nm particles, which were subsequently visualized by immunoelectron microscopy (data not shown). The supernatant from clone 1c was analyzed in a similar manner; no HBsAg activity was found anywhere in the gradient.

DISCUSSION

Amongst the few viruses implicated in human cancers, notably herpes simplex virus type 2, Epstein-Barr virus, human T-cell leukemia virus, and HBV, only in the case of HBV are viral sequences commonly found integrated into the genome of related tumor cells. We have cloned and determined the organization of integrated HBV sequences from a HCC and a HCC-derived cell line. Clone λ IA22 (derived from the liver tumor DNA) carries a complete genome plus an additional 25% of a HBV genome length. By contrast, clone λ A3 (derived from the cell line DNA) is made up of three subgenomic fragments that are 67%, 66%, and 11% of a complete genome. The two large subgenomic fragments are inverted with respect to one another. They possess the same restriction map and are probably derived from the same variant. The restriction maps of the two flanking cellular sequences of clone λ A3 are not the same, showing that this structure of HBV sequences probably did not arise by an amplification and reintegration involving the flanking cellular sequences. It is possible that the organization of HBV sequences in the two clones may not represent those after the initial integration event [it is estimated that the onset of HCC occurs between 15 and 25 years after HBV infection (24)]. This is particularly pertinent now that the presence of subgenomic fragments and rearranged sequences has been established. However, the finding of a complete integrated HBV genome shows that ordered integration does occur, and thus it is possible such structures may play a role in the life cycle of the virus. The restriction maps of the integrated HBV sequences in the two clones have been realigned with respect to the sequence of a cloned HBV genome [subtype ayw (22)] (Fig. 5). It is clear that the restriction maps of the two clones differ from each other as they do from all known HBV restriction maps (21). This rules out the simple hypothesis that there is a unique variant of HBV capable of integration into hepatocellular DNA.

For the two clones, host-viral junctions were located be-

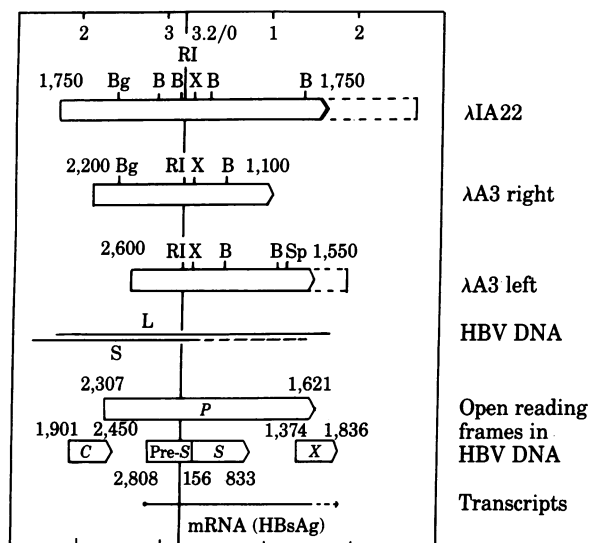


FIG. 5. Alignment of the restriction maps of the HBV sequences carried by cloned λ IA22 and λ A3 (divided into two) with respect to the common *EcoRI* site (RI; at positions 0 and 3.2 kb in free viral DNA). The approximate sites of HBV integration are given in bp. Restriction sites are codified as in Fig. 1. Also given are the long (L) and short (S) DNA strands and the open reading frames P, C, pre-S, S, and X in HBV DNA. The size of the known HBsAg mRNA is also shown.

tween 1,750 and 2,600 bp on the free viral genome. Edman et al. have found by a Southern blotting analysis of the cellular DNA with isolated HBV fragments as probes that, for the PLC/PRF/5 cell line, the hepatocyte-HBV junctions could be mapped to a similar region (1,400–2,600 bp) on the viral genome (25). HBV and WHV belong to the recently described hepadna group of viruses (26). Cloned integrated forms of WHV from woodchuck HCC have been reported (27). In two cases out of four, host-virus junctions mapped to about 2,000–2,100 bp on the viral genome. In conjunction with our results this suggests that the region between 1,750 and 2,600 bp, which carries the cohesive ends and coding portion of gene C, may be involved in the integration or reorganization of the DNA of HBV and homologous viruses. With respect to the hepatocyte genome we did not find any flanking sequences in common with those of the left side of clone λ IA22 amongst the 14 other HBV-containing phage clones. This supports the idea that there do not exist a few dispersed copies of a specific region of DNA linked to the sites of integrated HBV DNA.

Between clones λ IA22 and λ A3, only the entire pre-S/S open reading frame is conserved. By contrast, for integrated WHV sequences only the open reading frame X was conserved between the two clones reported (27). Our results do not exclude the possibility that region X is present in integrated HBV sequences because we have analyzed only one of a few HBV integration sites for both tumor and PLC/PRF/5 DNA.

In the hepatocytes of poorly differentiated HCC a common finding is the absence of HBsAg as determined by direct immunofluorescence (7). For patient 53, no HBsAg was found in the neoplastic hepatocytes (1). The trivial explanation that the corresponding HBV sequences are deleted or rearranged is clearly not the case, at least for one sequence in patient 53. In this case, it is likely that the neoplastic hepatocytes possessed a structurally intact but functionally inactive gene for HBsAg. Why HBsAg expression should be absent from HCC cells while the corresponding gene is intact is not understood. However, this finding parallels the case of simian virus 40 tumor antigen expression in murine teratocarcinoma-derived stem cells (28).

The recent discovery of a reverse transcriptase activity in the

duck hepatitis virus (29), a member of the HBV-like family of viruses, has prompted comparisons between these HBV-like viruses and the RNA retroviruses (30). This is particularly interesting in that, whereas HBV infection has been associated with the subsequent development of HCC (24), many of the RNA retroviruses are oncogenic. Amongst the 15 clones isolated we did not find any cellular homologues to six retroviral oncogenes in the sequences flanking the integrated HBV DNA. Because not all oncogenes are derived from retroviruses, transfection studies will be necessary to help in elucidating the possible role of integrated HBV DNAs and their cellular flanking sequences in hepatocyte transformation.

We thank Drs. Patrick Charnay, Alex Fritsch, Daniëlle Geldwerth, and John Collins for their comments and contribution to the work. We also thank Louise-Marie Da for preparing the manuscript. This work was supported by grants from the Institut National de la Santé et de la Recherche Médicale and the Faculté de Médecine Lariboisière Saint-Louis, Université de Paris VII.

- Bréchet, C., Pourcel, C., Louise, A., Rain, B. & Tiollais, P. (1980) *Nature (London)* **286**, 533–535.
- Bréchet, C., Hadchouel, M., Scotto, J., Fonck, M., Potet, F., Vyas, G. N. & Tiollais, P. (1981) *Proc. Natl. Acad. Sci. USA* **73**, 3906–3910.
- Bréchet, C., Hadchouel, M., Scotto, J., Degos, F., Charnay, P., Trépo, C. & Tiollais, P. (1981) *Lancet* **ii**, 765–768.
- Shafritz, D. A. & Kew, M. C. (1981) *Hepatology* **1**, 1–8.
- Shafritz, D. A., Shouval, D., Sherman, H. I., Hadziyannis, S. J. & Kew, M. C. (1981) *N. Engl. J. Med.* **305**, 1067–1073.
- Dejean, A., Vitvitski, L., Bréchet, C., Trépo, C., Tiollais, P. & Charnay, P. (1982) *Virology* **121**, 195–199.
- Nazarewicz, T., Krowczynski, K., Slusarczyk, J. & Nowoslawski, A. (1977) *J. Infect. Dis.* **135**, 298–302.
- Macnab, G. M., Alexander, J. J., Lecatsas, G., Bey, E. M. & Urbanowicz, J. M. (1976) *Br. J. Cancer* **34**, 509–515.
- Loenen, W. A. M. & Brammar, W. J. (1980) *Gene* **10**, 249–259.
- Pourcel, C., Marchal, C., Louise, A., Fritsch, A. & Tiollais, P. (1979) *Mol. Gen. Genet.* **170**, 161–169.
- Hohn, B. (1979) *Methods Enzymol.* **68**, 299–309.
- Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
- Wahl, G. M., Stern, M. & Stark, G. R. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3683–3687.
- Weinstock, R., Sweet, R., Weiss, M., Cedar, H. & Axel, R. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1299–1303.
- Ish-Horowitz, D. & Burke, J. F. (1981) *Nucleic Acids Res.* **9**, 2989–2998.
- Benton, W. D. & Davies, R. W. (1977) *Science* **196**, 180–182.
- Colbère-Garapin, F., Chousterman, S., Horodniceanu, F., Kourilsky, P. & Garapin, A. C. (1979) *Proc. Natl. Acad. Sci. USA* **78**, 3755–3759.
- Stow, N. D. & Wilkie, N. M. (1976) *J. Gen. Virol.* **33**, 447–458.
- Pourcel, C., Louise, A., Gervais, M., Chenciner, N., Dubois, M. F. & Tiollais, P. (1982) *J. Virol.* **42**, 100–105.
- Davis, R. W., Simon, M. & Davidson, N. (1971) *Methods Enzymol.* **21**, 413–428.
- Wain-Hobson, S., Pourcel, C. & Tiollais, P. (1982) in *Genetic Maps*, ed. O'Brien, S. J. (National Institutes of Health, Bethesda, MD), Vol. 2, pp. 57–64.
- Tiollais, P., Charnay, P. & Vyas, G. N. (1980) *Science* **213**, 406–411.
- Dejean, A., Carloni, G., Bréchet, C., Tiollais, P. & Wain-Hobson, S. (1982) *J. Cell. Biochem.*, in press.
- Szmuness, W. (1978) *Prog. Med. Virol.* **24**, 40–69.
- Edman, J. C., Gray, P., Valenzuela, P., Rall, L. B. & Rutter, W. (1980) *Nature (London)* **288**, 535–538.
- Summers, J., Smolec, M. J. & Snyder, R. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4533–4537.
- Ogston, C. W., Jonak, G. J., Rogler, C. E., Astrin, S. M. & Summers, J. (1982) *Cell* **29**, 385–394.
- Knowles, B. B., Pan, S., Solter, D., Linnenbach, A., Croce, C. & Huebner, K. (1980) *Nature (London)* **288**, 615–616.
- Summers, J. & Mason, W. S. (1982) *Cell* **29**, 403–415.
- Summers, J. & Mason, W. S. (1982) *Hepatology (N.Y.)*, Suppl. 2, 61S–66S.